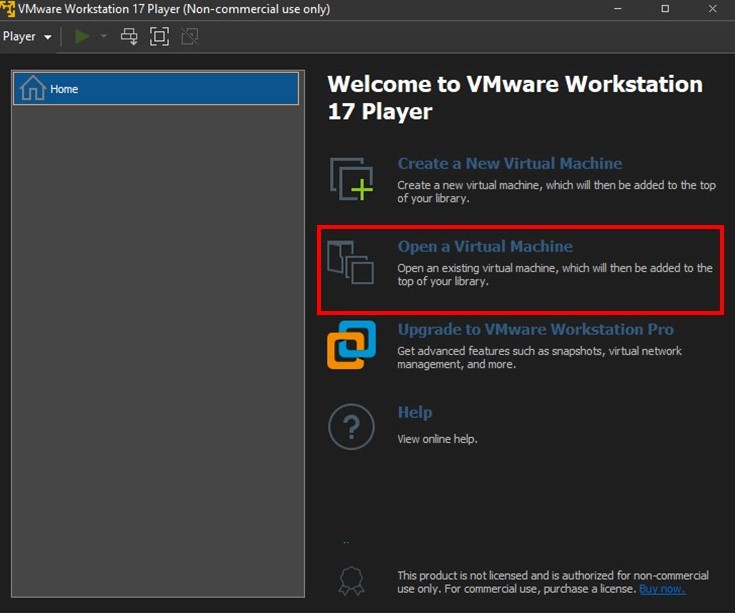
Step 1: Firstly, install VMWare Workstation using the files provided in prerequisites

folder. Then, download the Cloudera extension from the same folder and

extract its contents.

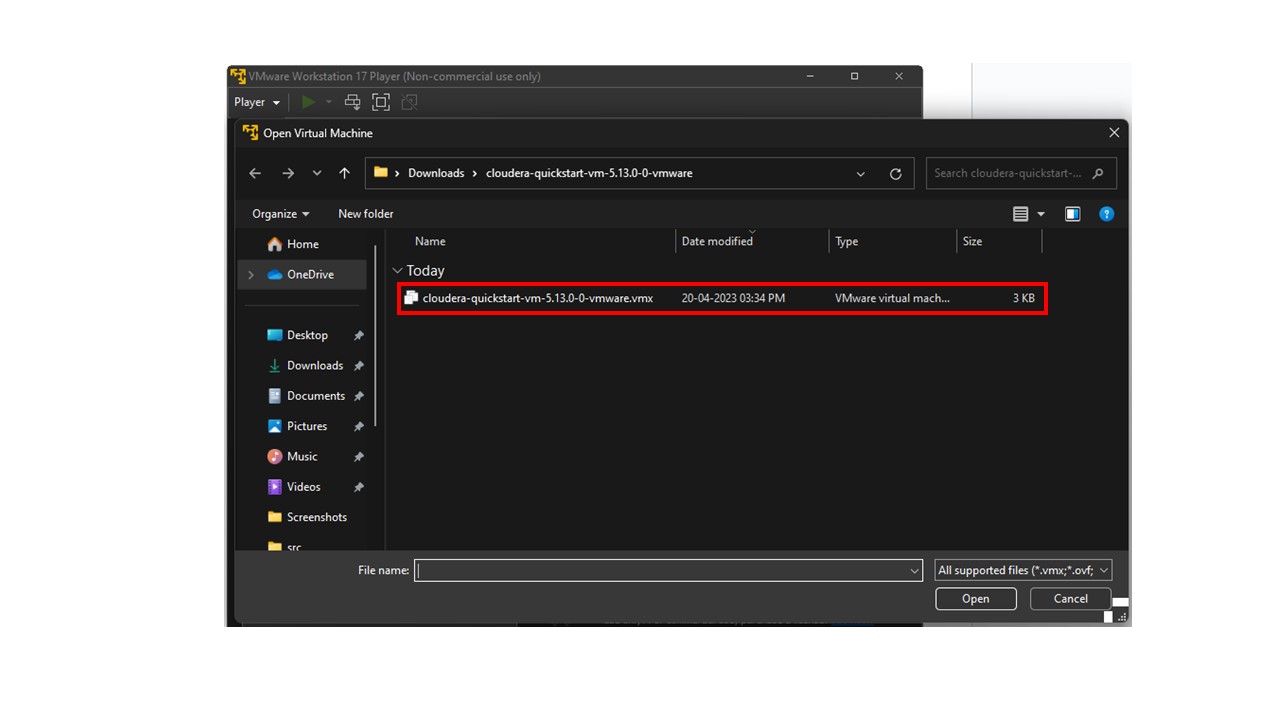
Step 2: After completing the previous step, locate **Open a Virtual Machine** and click

on the option to open it."

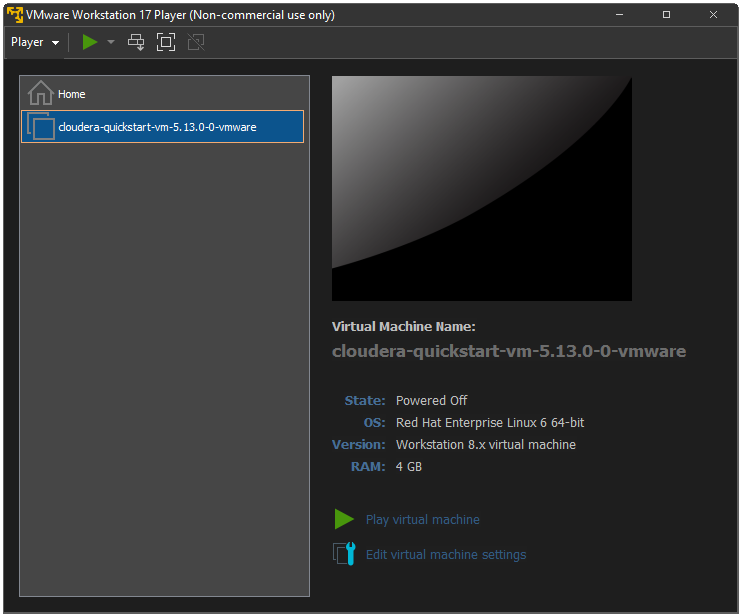


Step 3: To open the **Cloudera VM**, you need to locate the **Cloudera VM** file in your

directory and then access it



Step 4: Run Cloudera VM



Step 5: Create a folder with name weather analysis in a virtual machine by using following

Command.

**$ mkdir logfile\_analysis**

Step 6: Using the following command, navigate to the directory of the name weather

analysis in a virtual machine.

**$ cd logfile\_analysis/**

Step 7: Write programs for weather\_analysis to find the average for temperature, dew point, and wind speed using the given command, and utilize a text file to count the number of occurrences of each word.

**$ gedit LogFileMapper.java**

**$ gedit LogFileReducer.java**

**$ gedit LogFileCountryDriver.java**

**$ gedit log\_file.txt**

Step 8: Create a Hadoop directory for storing the TXT file and then put the TXT file

into that directory by using the given command.

**$ hdfs dfs -mkdir /logfile\_analysis/**

**$ hdfs dfs -put sample\_weather.txt /Weather\_analysis/**

Step 9: Compile the Java files using the following commands, which include the

Hadoop libraries.

**$ javac -cp $(hadoop classpath) \*.java**

Step 10:Create a Java archive for distribution using the following command

**$ jar -cvf weather.jar \*.class**

Step 11:To run a JAR file in Hadoop and see the average for temperature, dew

point, and wind speed in the TXT file, type the following command:

**$ hadoop jar weather.jar weather\_analysis /Weather\_analysis/sample\_**

**weather.txt /op**